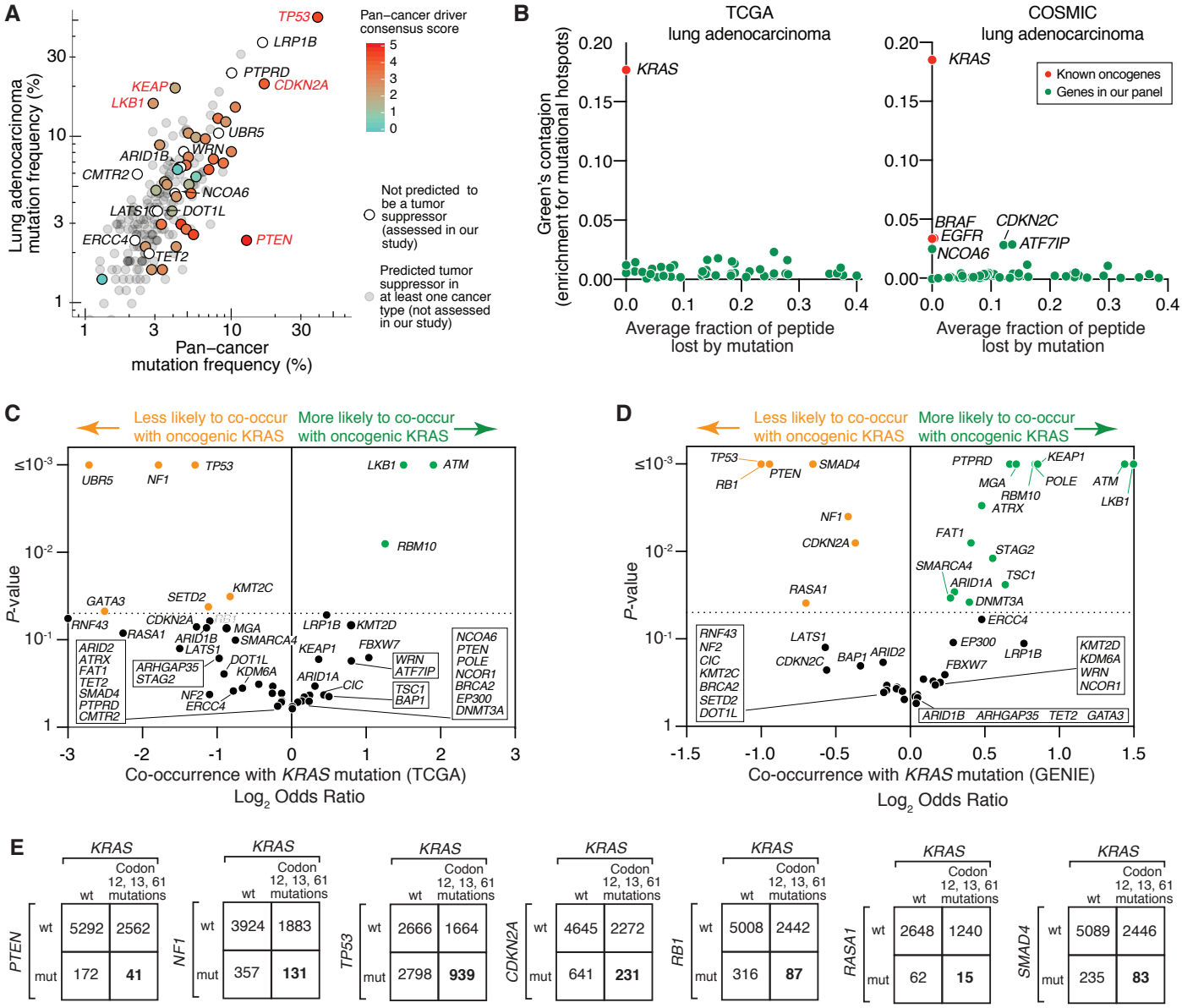


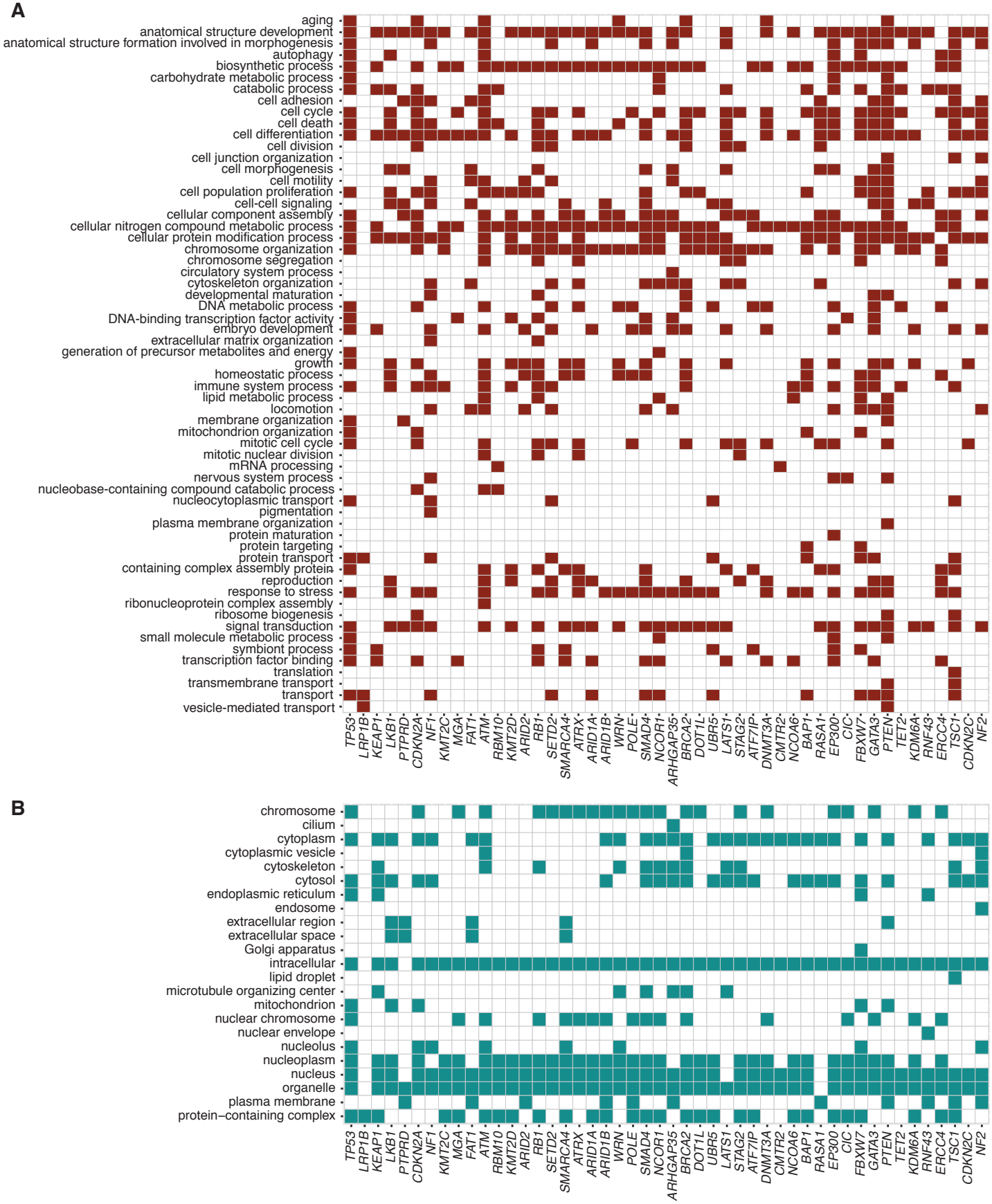
Supplementary Figure 1. Additional mutational characteristics of the putative tumor suppressor genes included in our analysis.

(A) Mutation frequencies across pan-cancer and in lung adenocarcinoma (data from TCGA). Color denotes pan-cancer driver consensus score derived from multiple prediction tools. Dots with black outline indicate genes that were included in our analysis, while gray dots represent other putative tumor suppressors that were not assessed. Several genes with high consensus scores are labelled in red. Genes labeled in black are those that we assessed but for which the pan-cancer consensus does not suggest tumor suppressor function.

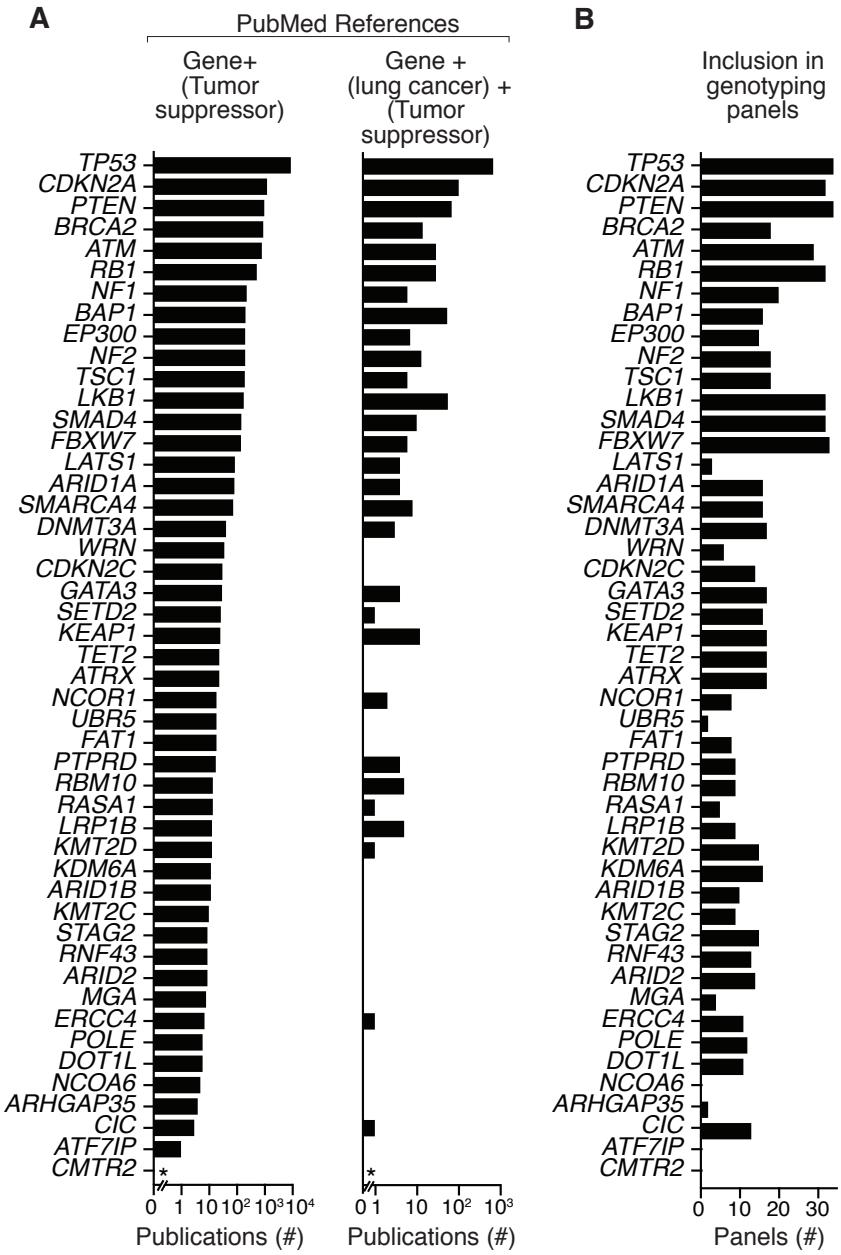
(B) Features of the mutations in each gene are consistent with tumor suppressor function. Green's contagion is a measure of mutational hotspots. Larger values indicate that mutations are enriched in particular regions of the protein, which is a typical feature for oncogenes. This measure of overdispersion is normalized to not scale with sample size and to be zero when mutations are randomly scattered across the transcript. Average fraction of protein lost by mutation combines that nonsense/frameshift mutation rate and location in the transcript of the mutations in each gene [(percent of protein transcript altering mutations that are nonsense or frameshift)*(Average fraction of protein lost by nonsense or frameshift mutations)]. TCGA plot does not include *CDKN2C*.

(C,D) Co-occurrence of mutations in KRAS and each putative tumor suppressor in TCGA PanCancer Atlas (566 lung adenocarcinoma cases) (C) and GENIE(8522 lung adenocarcinoma samples) (D). Dashed line indicates $P = 0.05$. Log_2 Odds Ratio and p value (one-sided Fisher's Exact Test) were both derived from cBioPortal.org.

(E) Despite being statistically mutually exclusive with oncogenic *KRAS*, *PTEN*, *NF1, TP53, CDKN2A, RB1, RASA1* and *SMAD4* mutations still occur concomitantly with oncogenic *KRAS* mutations in a subset of human lung adenocarcinoma, suggesting that the investigation of these genotypes in mouse models is still clinically relevant. The number of patients with each combination of genes mutated is listed (data from GENIE).

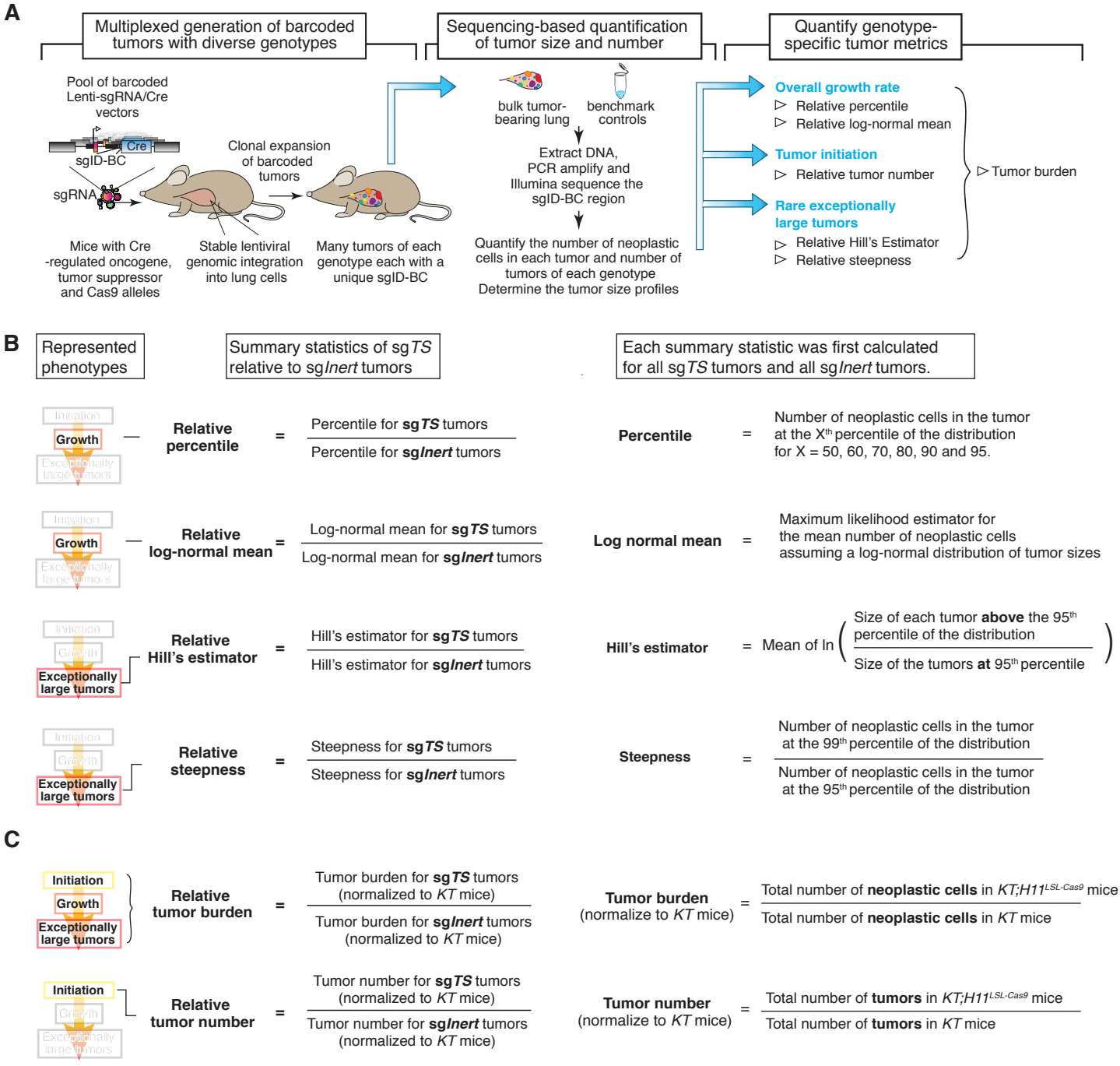


Supplementary Figure 2. The 48 putative tumor suppressor genes span a broad range of biological processes.
(A) Involvement of the 48 genes in different biological processes from Gene Ontology Slim (GO Slim) terms that provide a concise overview of the gene sets by using a subset of terms from gene ontologies (GO release date 2019-07-01, Ashburner *et al.* 2000, The Gene Ontology Consortium, 2019).
(B) The GO Slim annotated sub-cellular compartments in which the protein products of the 48 genes reside.



Supplementary Figure 3. Many putative tumor suppressors are not well studied and are not frequently included in cancer gene sequencing panels.

(A) To assess the extent to which each gene in our panel has been studied as a tumor suppressor in general, as well as in lung cancer, we ranked the genes by the number of publication accessible in PubMed with the indicated key words (Supplementary Methods). **CMTR2* has one paper that highlights it as a potential tumor suppressor gene in lung adenocarcinoma based on human mutation data, however this paper was only uncovered by manual annotation (Campbell *et al.*, Nature Genetics, 2016). (B) As an additional metric of the relative appreciation of each gene as a cancer driver, we determined the number of cancer gene sequencing panels that assess each gene. These 33 panels include FoundationOne CDx, Stanford Solid Tumor Actionable Mutation Panel (STAMP) and the 31 panels that contribute to GENIE.

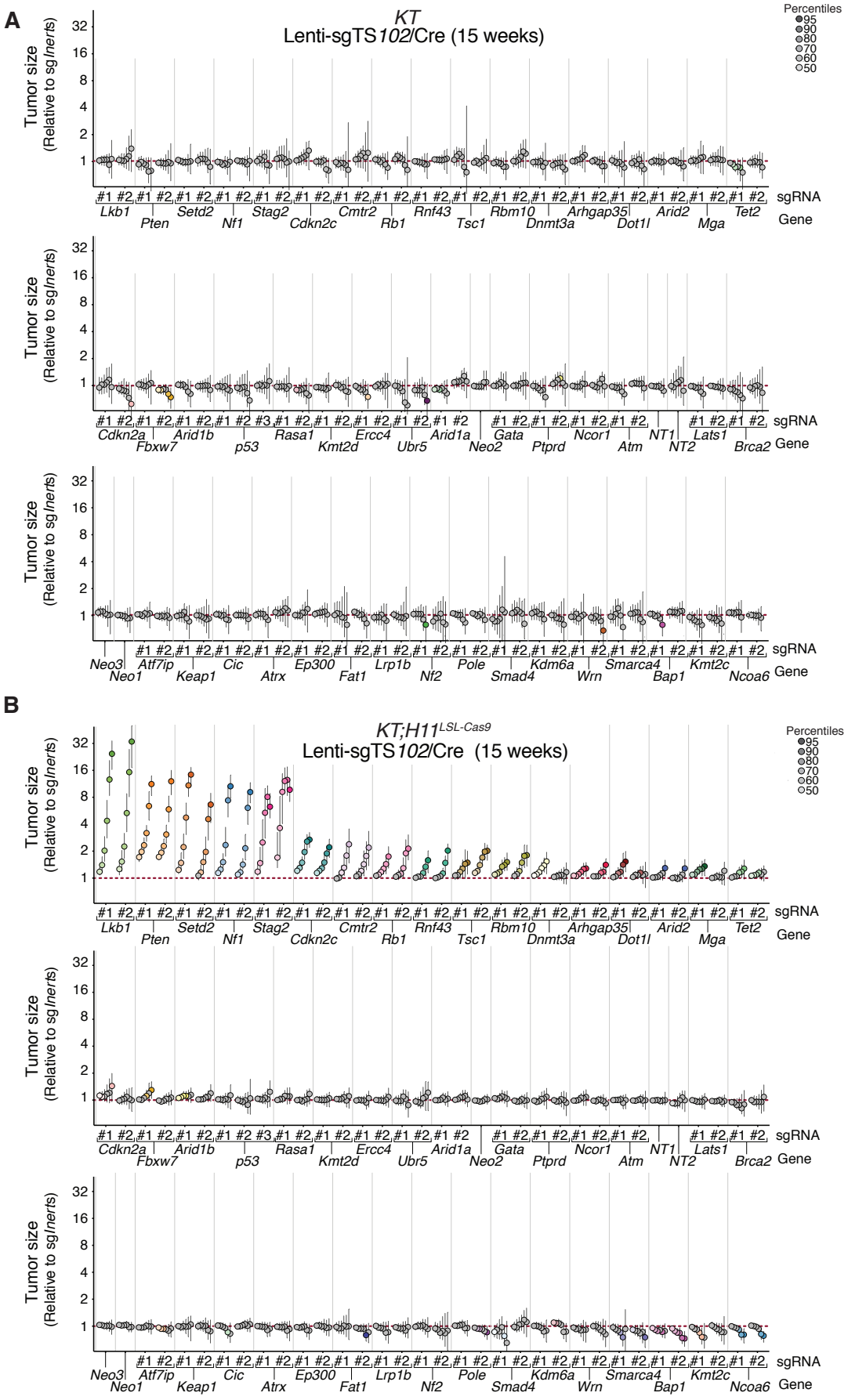


Supplementary Figure 4. Tumor phenotypes quantified by Tuba-seq can uncover the modes of action of tumor suppressor genes.

(A) Overview of tumor barcoding coupled with high-throughput barcode sequencing (Tuba-seq). The Tuba-seq pipeline includes three steps: Initiation of barcoded lung tumors with a pool of Lenti-sgRNA/Cre vectors, barcode sequencing-based analysis of tumor sizes, and statistical analyses to uncover different aspects of tumor suppressor function. The quantitative and multiplexed nature of Tuba-seq makes it capable of investigating tens and even hundreds of targets simultaneously to identify functional tumor suppressors. By generating each tumor genotype along with control tumors (with sg*Inert*) in the same mouse, this method minimizes the effect of mouse-to-mouse variability while enhancing the signal-to-noise ratio.

(B) Summary statistics that are based on taking the ratio of the statistic for tumors with a specific sgRNA targeting candidate tumor suppressor genes (sg*TS)* relative to that of sg*Inert* tumors. These statistics describe the advantage conferred by inactivation of the tumor suppressor gene. These summary statistics are not influenced by the representation of each vector in the pool and are calculated separately for *KT;H11*$^{LSL\text{-}Cas9}$ and *KT* mice.

(C) Summary statistics that are impacted by viral pooling are calculated for *KT;H11*$^{LSL\text{-}Cas9}$ mice and normalized to *KT* mice. The number of neoplastic cells and the number of tumors depend on the relative titer of each viral vector. Therefore, these metrics in *KT;H11*$^{LSL\text{-}Cas9}$ mice are normalized to that in *KT* mice with tumors initiated with the exact same pool of viral vectors to account for slight differences in the titer of individual Lenti-sgRNA/Cre vectors. In (B) and (C), relative summary statistics are listed on the left and the calculation of each summary statistics for sg*TS* or sg*Inert* tumors is shown on the right.



Supplementary Figure 5. Tumor sizes at defined percentiles within the distribution of tumors initiated with each Lenti-sgRNA/Cre vector.

Legend on next page

Supplementary Figure 5. Tumor sizes at defined percentiles within the distribution of tumors initiated with each Lenti-sgRNA/Cre vector.

(A,B) Tumor size (neoplastic cell number) at the indicated percentiles for each barcoded Lenti-sgRNA/Cre vector relative to the average of sg*Inert*-containing tumors in *KT* mice (A) and *KT;H11*$^{LSL\text{-}Cas9}$ mice (B). Note that the first 17 genes in panel (B) are also shown in Fig. 2B. Genes are ordered based on the average of the 95th percentile tumor sizes from all sgRNAs targeting that gene in *KT;H11*$^{LSL\text{-}Cas9}$ (same as in Fig. 2A). Error bars indicate 95% confidence intervals. Percentiles that are significantly different from average of sg*Inert*s are in color. The 95% confidence intervals and *P*-values were calculated by bootstrap.

Plots represent aggregated data from 12 *KT* and 47 *KT;H11*$^{LSL\text{-}Cas9}$ mice (see Fig. 1C).

A

TCGA Lung Adenocarcinomas (n=507)

	missense mutations	homozygous deletions	translocations	truncating mutations	Total
STAG2	16	0	2	3	21
RAD21	6	0	0	2	8
SMC1A	7	2	1	0	10
SMC3	13	2	0	0	15
					51/507 (10.1%)

B

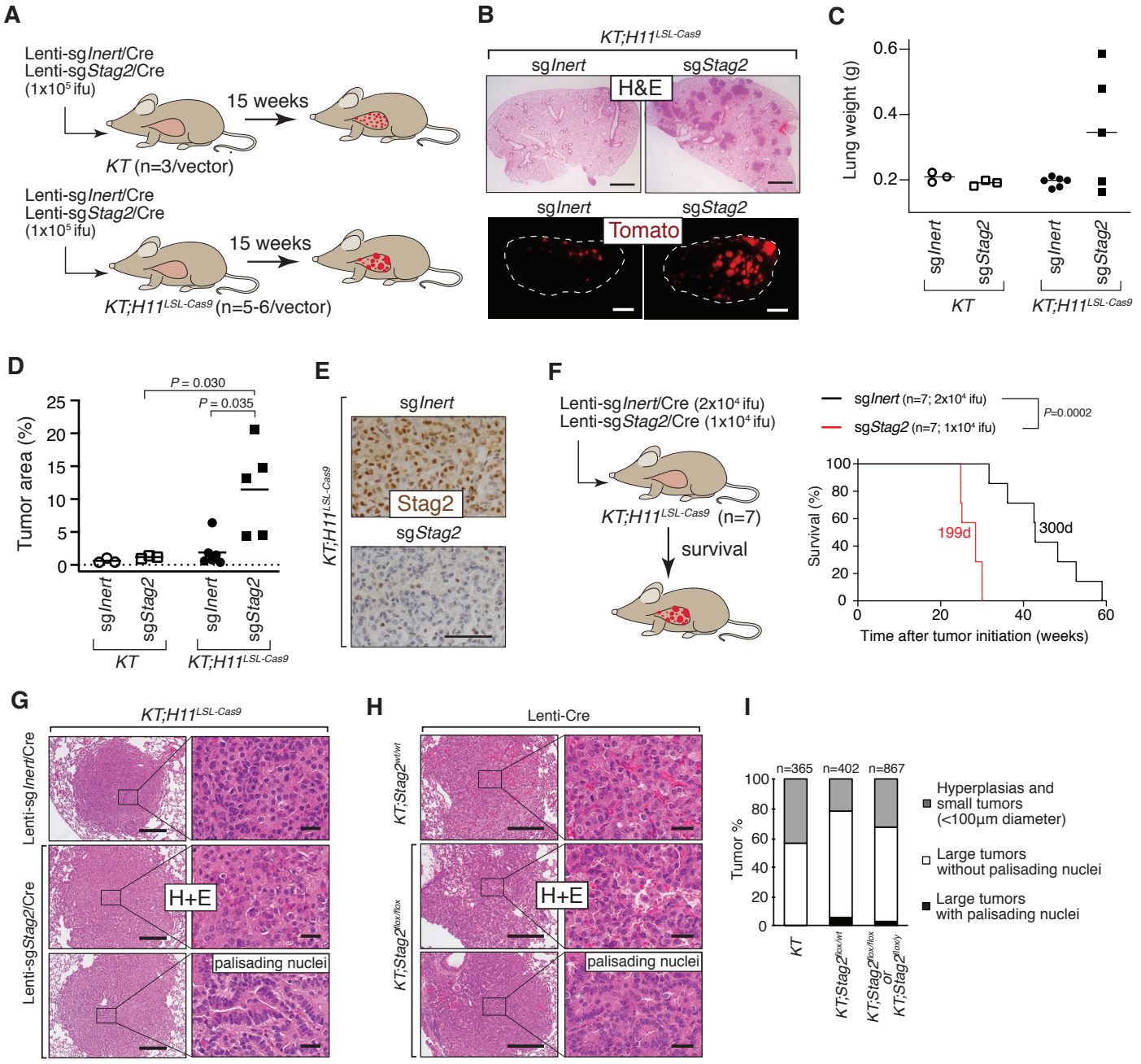
TRACERx Lung Adenocarcinomas (n=61)

	clonal mutations	subclonal mutations	Total
STAG2	0	2	2
RAD21	2	0	2
SMC1A	0	4	4
SMC3	0	2	2
			10/61 (16.4%)

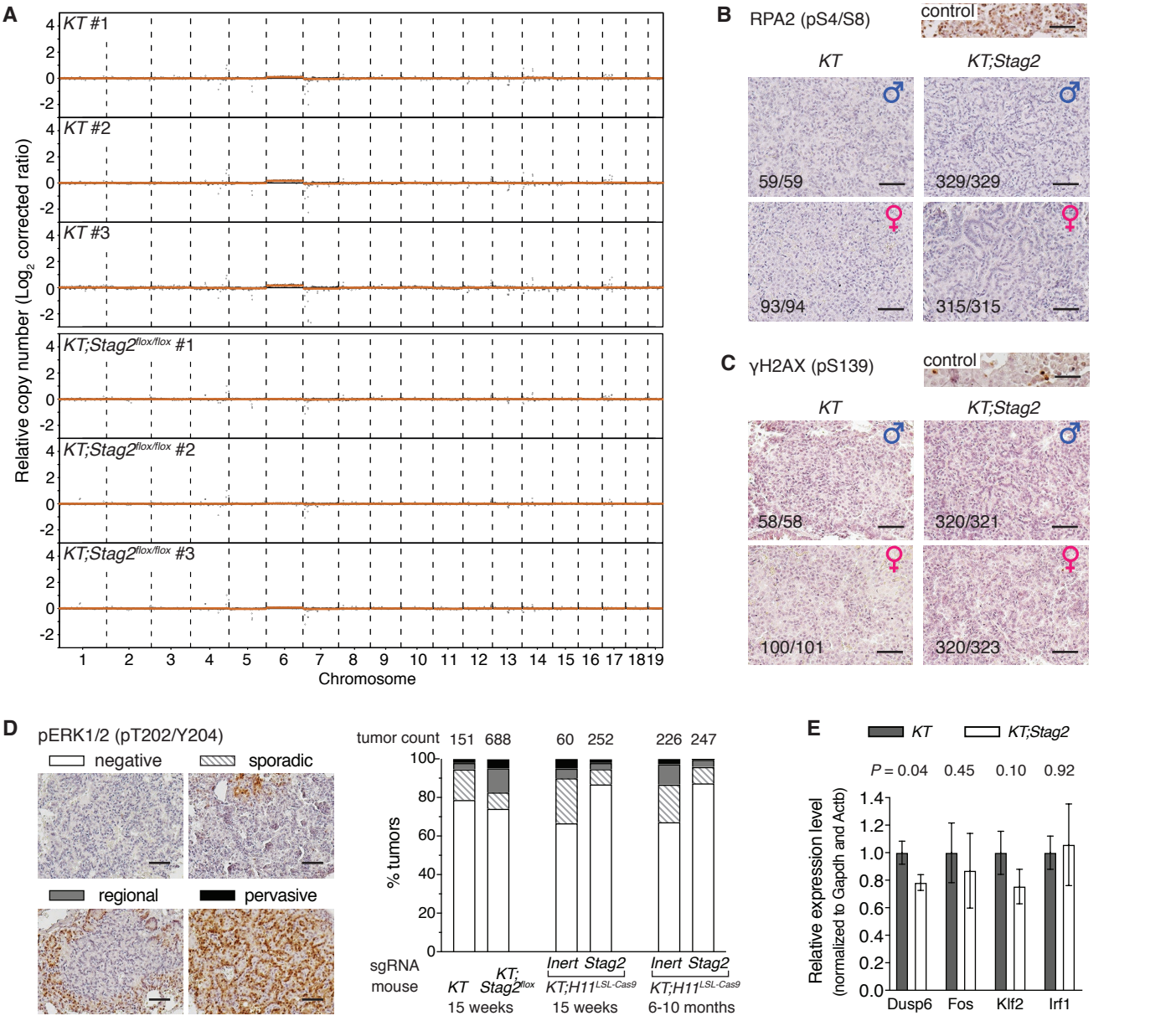
Supplementary Figure 6. Mutation frequency of *STAG2* and other members of the cohesin complex in human lung adenocarcinoma.

(A) Number of lung adenocarcinomas with mutations in the indicated genes (data from TCGA). The type of mutations is shown. Total number (and percent) of samples with one or more of these genes mutated is shown.

(B) Number of lung adenocarcinomas with mutations in the indicated genes (data from TRACERx). Whether the mutations were clonal or subclonal across multiple tumor regions is indicated. Total number (and percent) of samples with one or more of these genes mutated is shown.



Supplementary Figure 7. Analysis of Stag2-mediated tumor suppression in individual groups of mice.
(A) Schematic of tumor initiation with the indicated lentiviral vectors in *KT* and *KT;H11*$^{LSL\text{-}Cas9}$ mice. Mouse genotype, mouse number and titer of virus are indicated.
(B) H&E stained sections and fluorescence images of the lungs from *KT;H11*$^{LSL\text{-}Cas9}$ mice 15 weeks after tumor initiation with the indicated lentiviral vectors. Scale bars = 1 mm.
(C) Weight of the lungs from *KT* and *KT;H11*$^{LSL\text{-}Cas9}$ mice 15 weeks after tumor initiation with the indicated lentiviral vectors. Each dot represents a mouse and the bar is the mean.
(D) Quantification of tumor area (%) (tumor area/total lung area x 100) on H&E-stained sections of mouse lungs 15 weeks after tumor initiation. Each dot represents a mouse and the bar is the mean. *P*-values were calculated by Student's t-test.
(E) Lenti-sg*Stag2*/Cre initiated tumors in *KT;H11*$^{LSL\text{-}Cas9}$ mice lack Stag2 protein. Scale bar = 50 μm.
(F) Survival curve of *KT;H11*$^{LSL\text{-}Cas9}$ mice with tumors initiated with the indicated viral vectors. Mouse number and lentiviral titer are indicated. *P*-values from Mantel-Haenszel test. Median survival (in days) is indicated.
(G) H&E stained sections of representative lung tumors from *KT;H11*$^{LSL\text{-}Cas9}$ mice 6 months after tumor initiation with the indicated lentiviral vectors. Left scale bars = 200 μm; right scale bars = 20 μm.
(H) Representative H&E images of lung tumor from the indicated genotypes of mice. Left scale bars = 200 μm; right scale bars = 20 μm.
(I) Quantification of lung tumors with and without palisading nuclei based on histology. The numbers of tumors quantified are shown above each column.



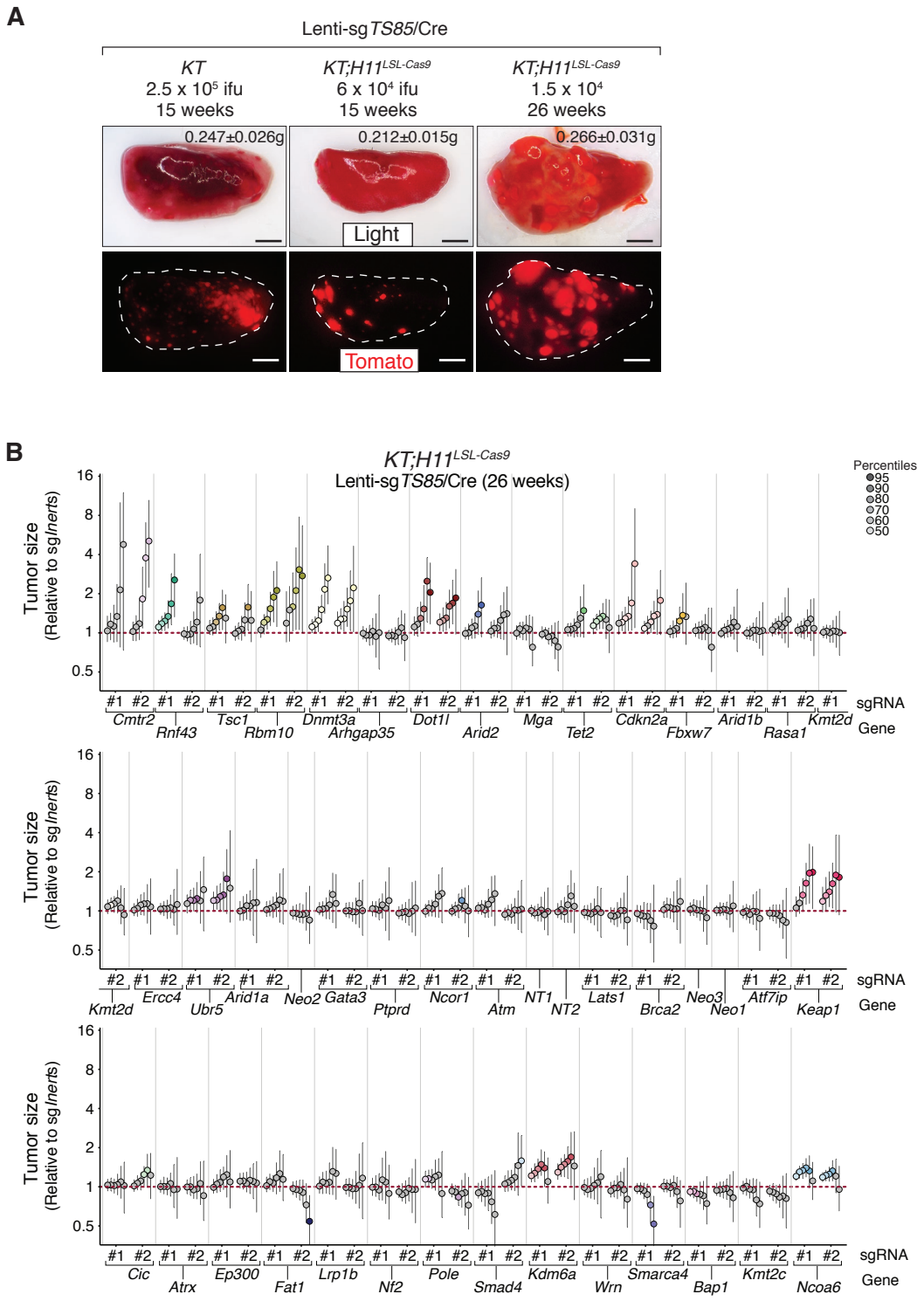
Supplementary Figure 8. *Stag2*-deficient lung tumors do not exhibit widespread aneuploidy, DNA damage response, Mek/Erk activation, or cGas/Sting pathway activation.

(A) Sorted neoplastic cells (60,000-100,000 Tomato$^+$;Lineage$^-$ cells) were subjected to 0.1x whole genome sequencing. The data were analyzed using QDNASeq as previously described (Scheinin, 2014). This analysis uses 1MB bins to fit the copy number profile of the samples. These data suggest no overt aneuploidy in *Stag2*-deficient or proficient tumors.

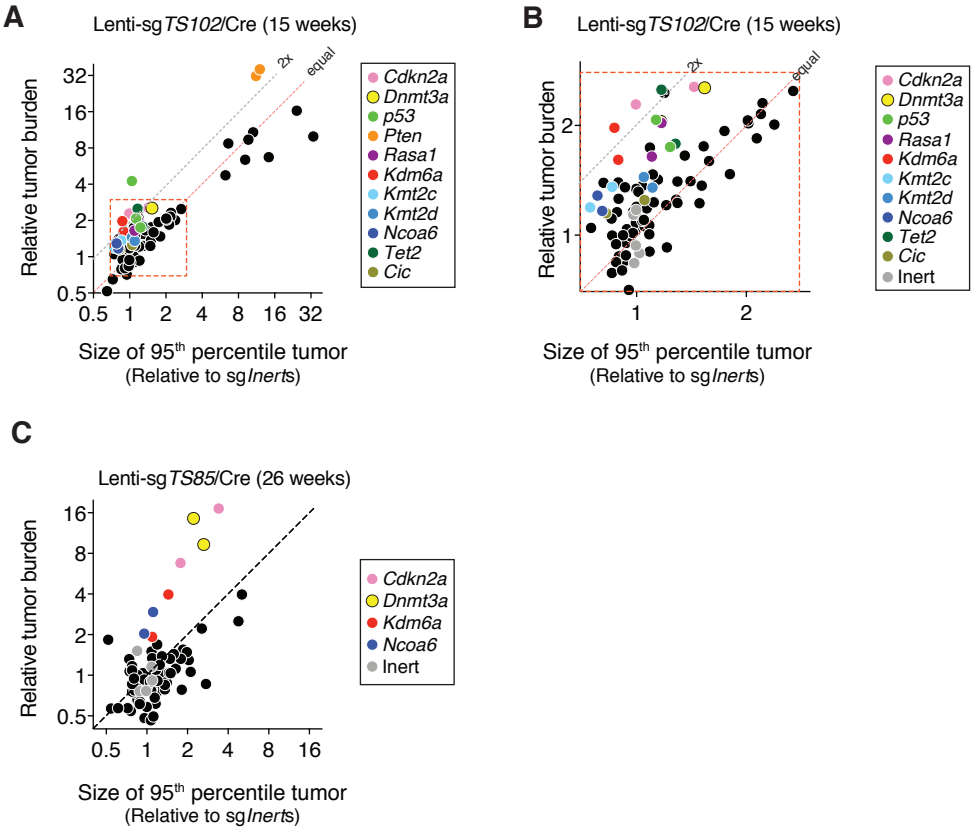
(B,C) Three male and three female mice per genotype (see Fig. 3A) were immunohistochemically stained for phospho-RPA2 (B) or phospho-histone H2A.X (C) and counterstained with hematoxylin. Control sections are from a murine SCLC. The number of phospho-RPA2- or γH2A.X-negative tumors / total tumors is indicated. Negative tumors were defined as tumors with ≤ 2 positive nuclei per 40x field. Scale bars = 50 μm. Most tumors were negative for these two markers, suggesting that Stag2 inactivation did not cause increased DNA damage.

(D) Stag2-deficient and proficient tumors were immunohistochemically stained for phospho-Erk1/2 and counterstained with hematoxylin. Representative images of tumors that are negative for pErk, with sporadic pErk^{positive} cells, with regional pErk^{positive} cells, or with pervasive pErk^{positive} cells are shown on the left. *KT* and *KT;Stag2^{flox}* samples (Fig. 3A) as well as sg*Inert* and sg*Stag2* samples (15 weeks, Supplementary Fig. S7A, and survival experiment sample, Supplementary Fig. S7F) were quantified. The percentage of tumors within each category are plotted, and the total number of quantified tumors are indicated above the columns. Quantification shows that Stag2 inactivation does not consistently increase Erk phosphorylation in all experiments, suggesting that Mek/Erk signaling is unlikely the major mechanism for Stag2-specific tumor suppression. Scale bars = 50 μm.

(E) The relative expression of *Dusp6*, Erk-target genes *Fos* and *Klf2*, as well as cGas/Sting target gene *Ifnl3, Ifnb1, Ifi44, Mx1* and *Irf1* were determined by quantitative real-time PCR. *P*-values indicated are calculated by Student's t-test. *Ifnl3*, *Ifnb1*, *Ifi44* and *Mx1* are not plotted because they were below the level of detection in all *KT* and *KT;Stag2* tumors even after 55 cycles. The other genes were not upregulated in *Stag2*-deficient tumors, suggesting that cGas/Sting pathway is unlikely driving the increased growth of *Stag2*-deficient tumors.



Supplementary Figure 9. Inactivation of some tumor suppressor genes generate greater effects at a later time point.
(A) Bright-field and fluorescence images of the lungs from the indicated groups of mice at the indicated time point after tumor initiation. Scale bars: 2 mm. Mean ± standard error of lung weight is indicated.
(B) Sizes of tumors at the indicated percentiles for each Lenti-sgRNA/Cre vector relative to that of sg*Inert*-targeted tumors in *KT;H11*$^{LSL\text{-}Cas9}$ mice. Error bars indicate 95% confidence intervals. Percentiles that are significantly different from the average of sg*Inert*s are in color. Genes are in the same order as in Fig. 2A.



Supplementary Figure 10. Inactivation of some tumor suppressor genes has a disproportionate impact on tumor burden relative to 95th percentile tumor size.

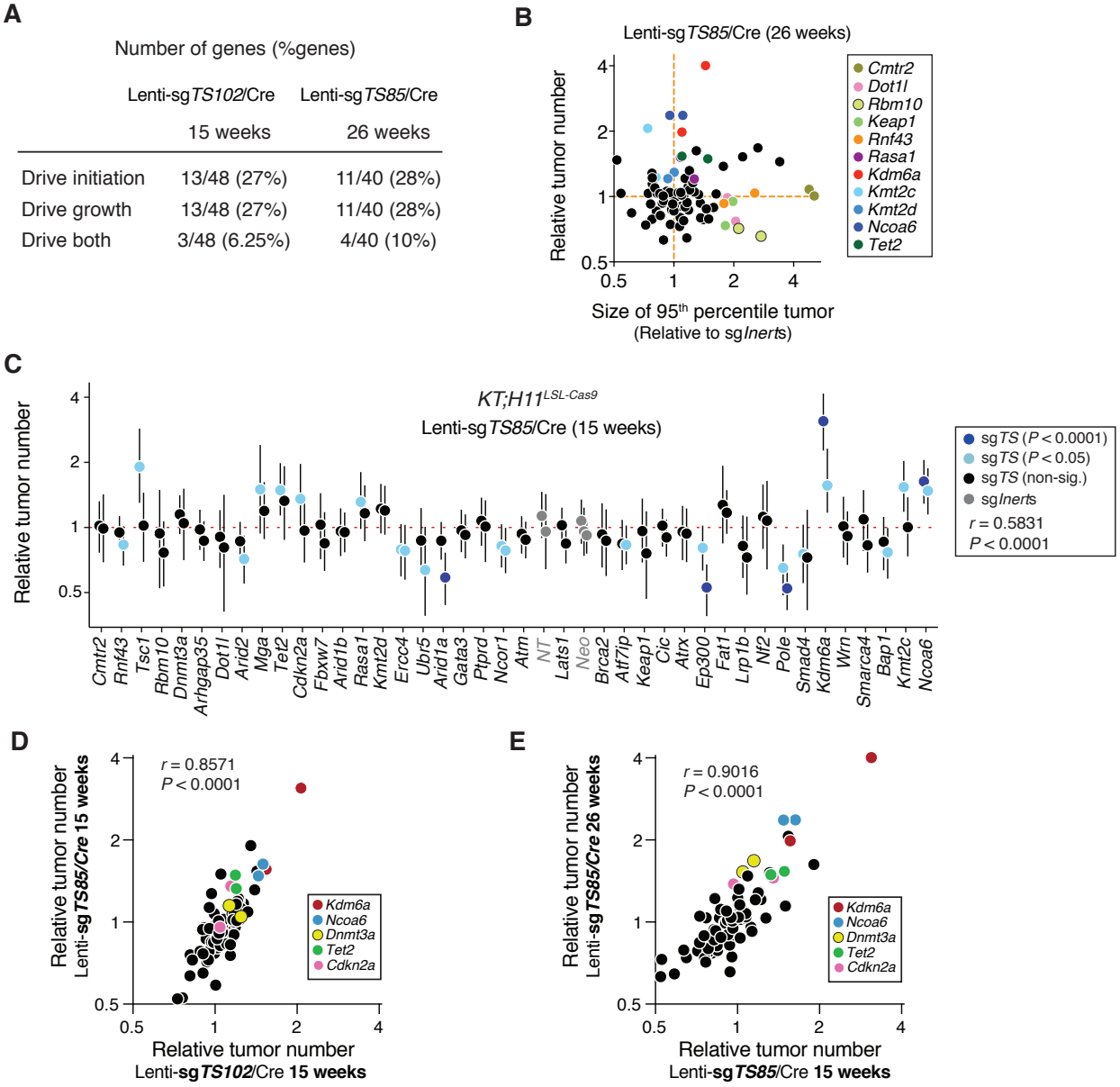
(A,B) Tumor suppressor effect on tumor burden and 95th percentile size (a metric of overall tumor growth) are generally concordant. However, some genes have a great effect of tumor burden suggesting that other aspects of tumor initiation and growth are controlled by those genes. Each dot represents an sgRNA. Data from tumors initiated with the Lenti-sg*TS102*/Cre in *KT;H11*$^{LSL\text{-}Cas9}$ mice are shown. sg*Inert*s are in gray and the sgRNA corresponding to the indicated genes are colored. Panel (A) shows all sgRNAs. Panel (B) has a truncated scale and shows the sgRNA that are in the boxed region in (A).

(C) Same plot as in (A) except for tumors initiated with the Lenti-sg*TS85*/Cre in *KT;H11*$^{LSL\text{-}Cas9}$ mice.

A

Number of genes (%genes)

	Lenti-sg*TS102*/Cre 15 weeks	Lenti-sg*TS85*/Cre 26 weeks
Drive initiation	13/48 (27%)	11/40 (28%)
Drive growth	13/48 (27%)	11/40 (28%)
Drive both	3/48 (6.25%)	4/40 (10%)



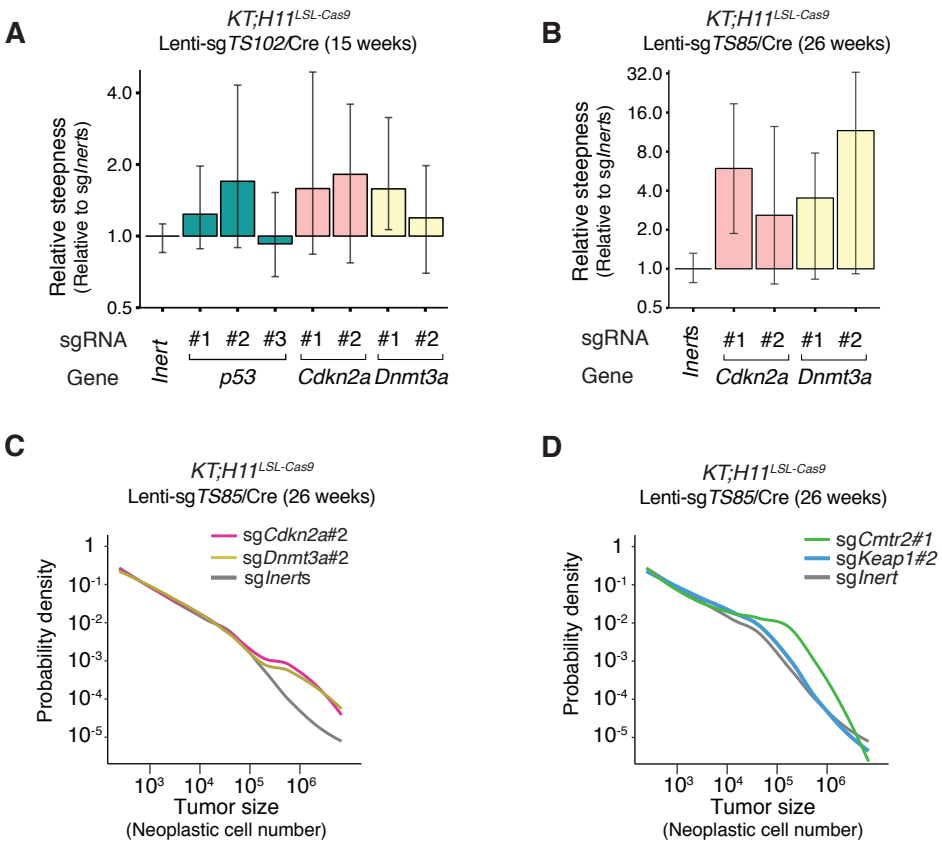
Supplementary Figure 11. The impact of tumor suppressor gene inactivation on tumor number is consistent across multiple independent experiments and multiple time points.

(A) The number of genes whose loss significantly drives tumor initiation, tumor growth or both. Data from tumors initiated with the Lenti-sg*TS102*/Cre for 15 weeks and the Lenti-sg*TS85*/Cre for 26 weeks are shown. When tumor initiation and growth are driven by independent machineries, the percentage of genes that drive both initiation and growth is expected to be 7.3% (15 weeks) and 7.8% (26 weeks). The actual percentage is close to expectation, indicating that gene sets that drive initiation and that drive growth are independent.

(B) Genotype specific effects on growth (represented by the size of the tumor at the 95th percentile) and tumor number can be independent aspects of tumor suppression. Data from tumors initiated with the Lenti-sg*TS85*/Cre in *KT;H11*$^{LSL\text{-}Cas9}$ mice for 26 weeks are shown.

(C) Inactivation of many tumor suppressor genes increases tumor number. Error bars indicate 95% confidence intervals. The effect of each sgRNA on tumor number 15 weeks after tumor initiation with Lenti-sg*TS85*/Cre in *KT;H11*$^{LSL\text{-}Cas9}$ mice is shown. sgRNAs that significantly change tumor number in *KT;H11*$^{LSL\text{-}Cas9}$ mice are colored as indicated. sg*Inert*s are in gray and dotted line indicates no effect. Genes are ordered as in Fig. 2A.

(D,E) Effects of tumor suppressor genes on tumor number are highly reproducible across all three Tuba-seq datasets. Lenti-sg*TS85*/Cre 15 weeks versus Lenti-sg*TS102*/Cre 15 weeks (D) and versus Lenti-sg*TS85*/Cre 26 weeks (E) are shown. Each dot represents an sgRNA. Pearson correlation coefficient (*r*, indicated in plots) shows correlation.



Supplementary Figure 12. Inactivation of *Cdkn2a* or *Dnmt3a* generates tumor size distributions consistent with the emergence of rare but exceptionally large tumors.

(A) As an alternative metric to quantify the emergence of rare but exceptionally large tumors, we calculated the ratio of the size of 99th percentile tumor to the 95th percentile tumor for each sgRNA and normalized this value to that of tumors with sg*Inert*s. Error bars are 95% confidence intervals. This metric termed "Relative steepness" suggests that *Dnmt3* and *Cdkn2a* inactivation allow the emergence of exceptionally large tumors at the tail of the tumor size distribution. Statistics are calculated from aggregating all tumors from 47 *KT;H11*$^{LSL\text{-}Cas9}$ mice.

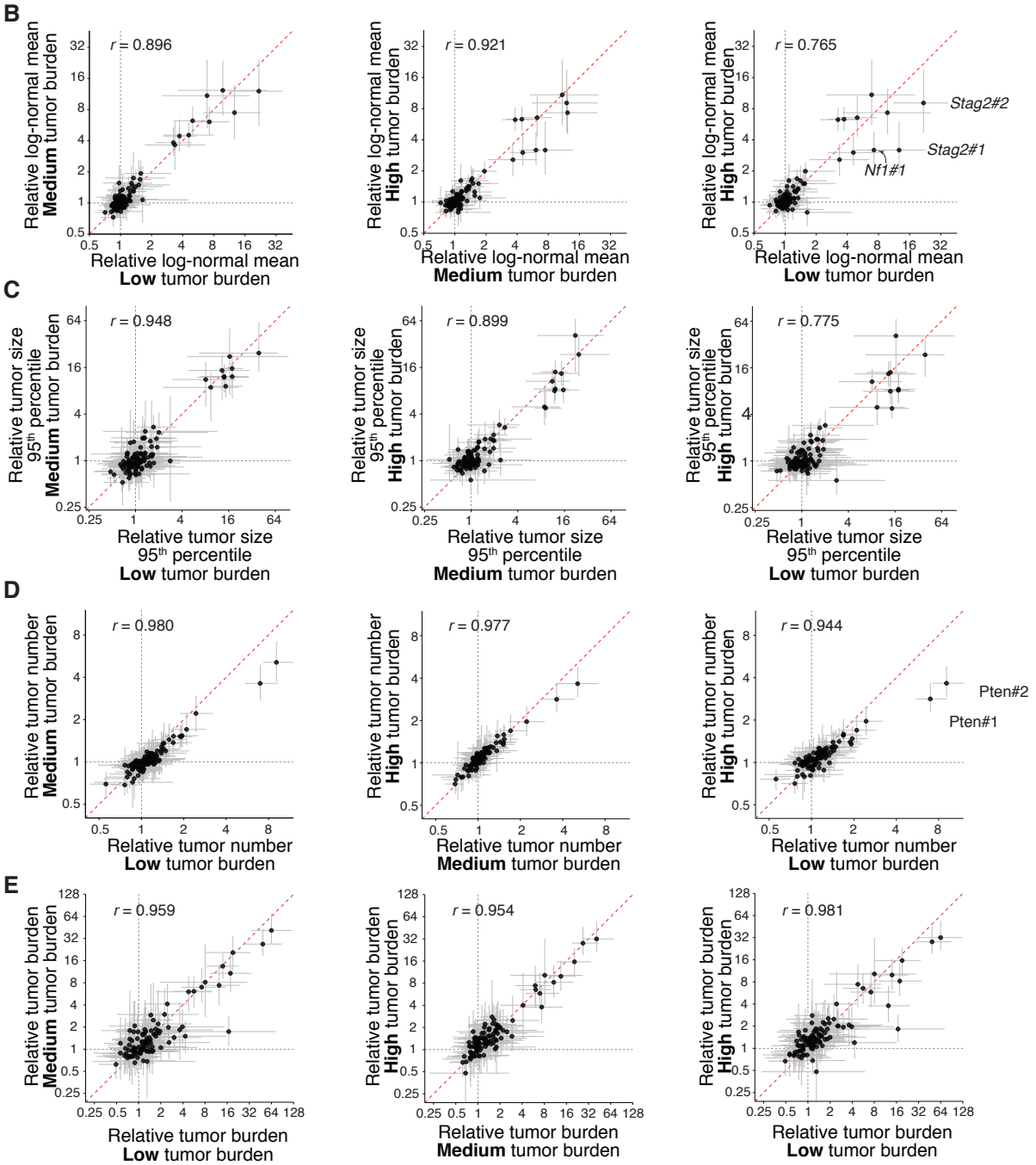
(B) Relative steepness supports that *Dnmt3* and *Cdkn2a* inactivation allow the emergence of exceptionally large tumors at the tail of the tumor size distribution. Error bars are 95% confidence intervals. Statistics are calculated from aggregating all tumors from 40 *KT;H11*$^{LSL\text{-}Cas9}$ (26 weeks) mice.

(C) Inactivation of *Dnmt3a* and *Cdkn2a* uniquely generate tumor size distributions with heavy tails. Probability density plots for tumors sizes show the profile of aggregated tumors with sg*Inert*s as well as the sgRNAs targeting either *Dnmt3a* or *Cdkn2a*. Data is aggregated from all tumors from 40 *KT;H11*$^{LSL\text{-}Cas9}$ (26 weeks) mice.

(D) In contrast, inactivation of *Cmtr2* or *Keap1* increases overall tumor growth but does not generate disproportionately large tumors. Probability density plot for tumors sizes show the profile of aggregated tumors with sg*Inert*s and representative profiles of *Cmtr2*- and *Keap1*-targeted tumors for comparison. Data is aggregated from all tumors from 40 *KT;H11*$^{LSL\text{-}Cas9}$ (26 weeks) mice.

A

	neoplastic cells (x10^6)	
Low tumor burden	5.35 +/– 2.20	16 mice
Medium tumor burden	14.57 +/– 4.76	16 mice
High tumor burden	42.23 +/– 18.67	15 mice



Supplementary Figure 13. Tumor initiation and growth metrics are largely uneffected by crowding effects due to high tumor burden.

(A) We divide the 47 *KT;H11*$^{LSL\text{-}Cas9}$ mice with Lenti-sg102/Cre initiated tumors into three groups; low tumor burden, medium tumor burden, and high tumor burden. The mean +/- SD of total neoplastic cell number and number of mice in each group is shown. Notes that on average low tumor burden and high tumor burden groups have almost a 10x difference in total neoplastic cell number.

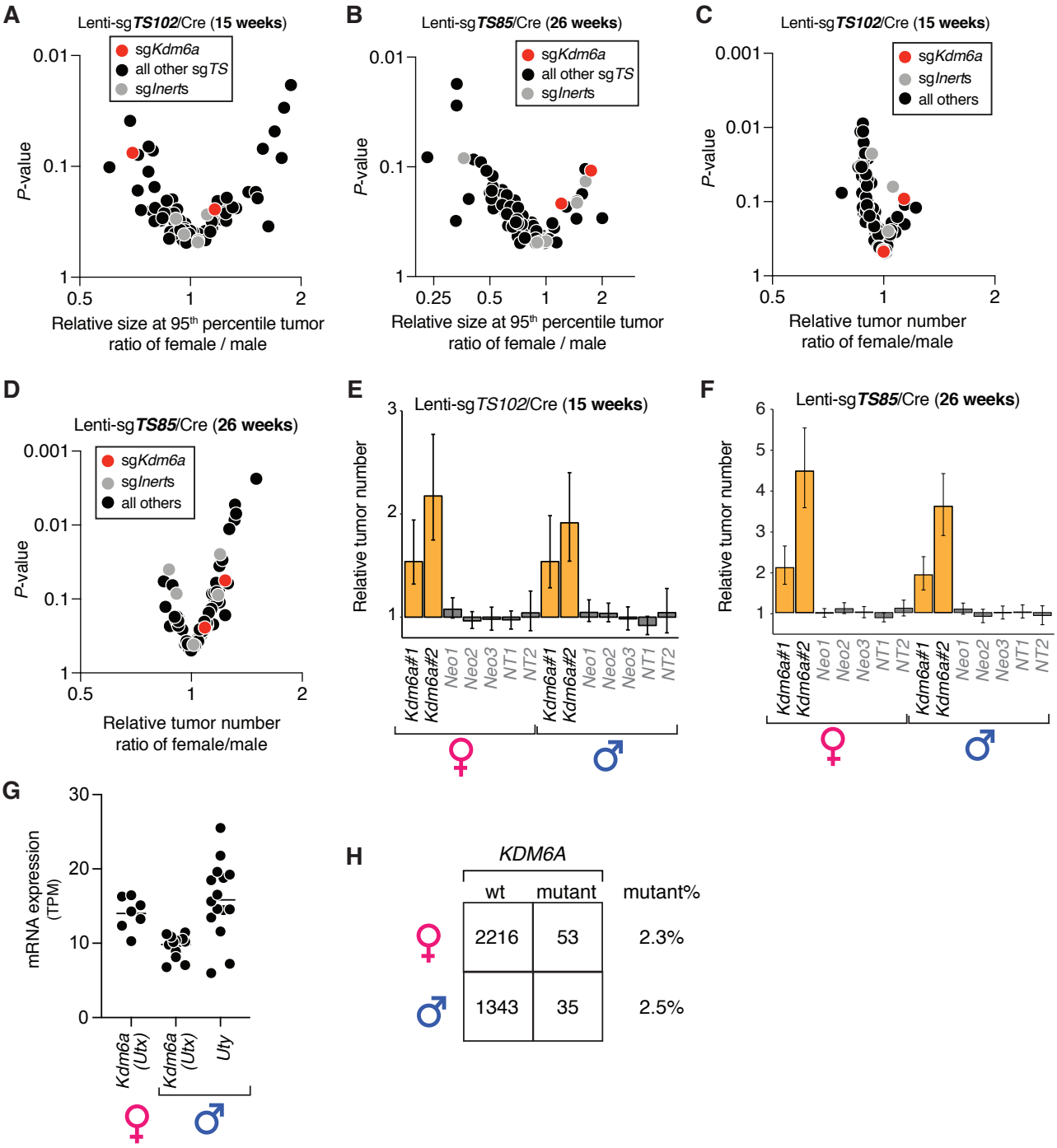
(B) The relative log-normal mean in each group plotted against those in each of the other groups. Three sgRNAs which deviate from the diagonal line are indicated.

(C) The relative tumor sizes at the 95th percentile in each group plotted against those in each of the other groups.

(D) The relative tumor number in each group plotted against those in each of the other groups. Two sgRNAs which deviate from the diagonal line are indicated.

(E) The relative tumor burden in each group plotted against those in each of the other groups.

In all plots, the black dots represent the estimate for each sgRNA and the gray horizontal and vertical bars represent the 95% confidence intervals. The red dashed lines represents the diagonal line. Pearson's correlation coefficients (*r*) are indicated in each plot. All estimated values are distributed around the diagonal line and are highly correlated.



Supplementary Figure 14. COMPASS complex components are frequently mutated in human lung adenocarcinoma with no indication of a gender-specific effect of Kdm6a.

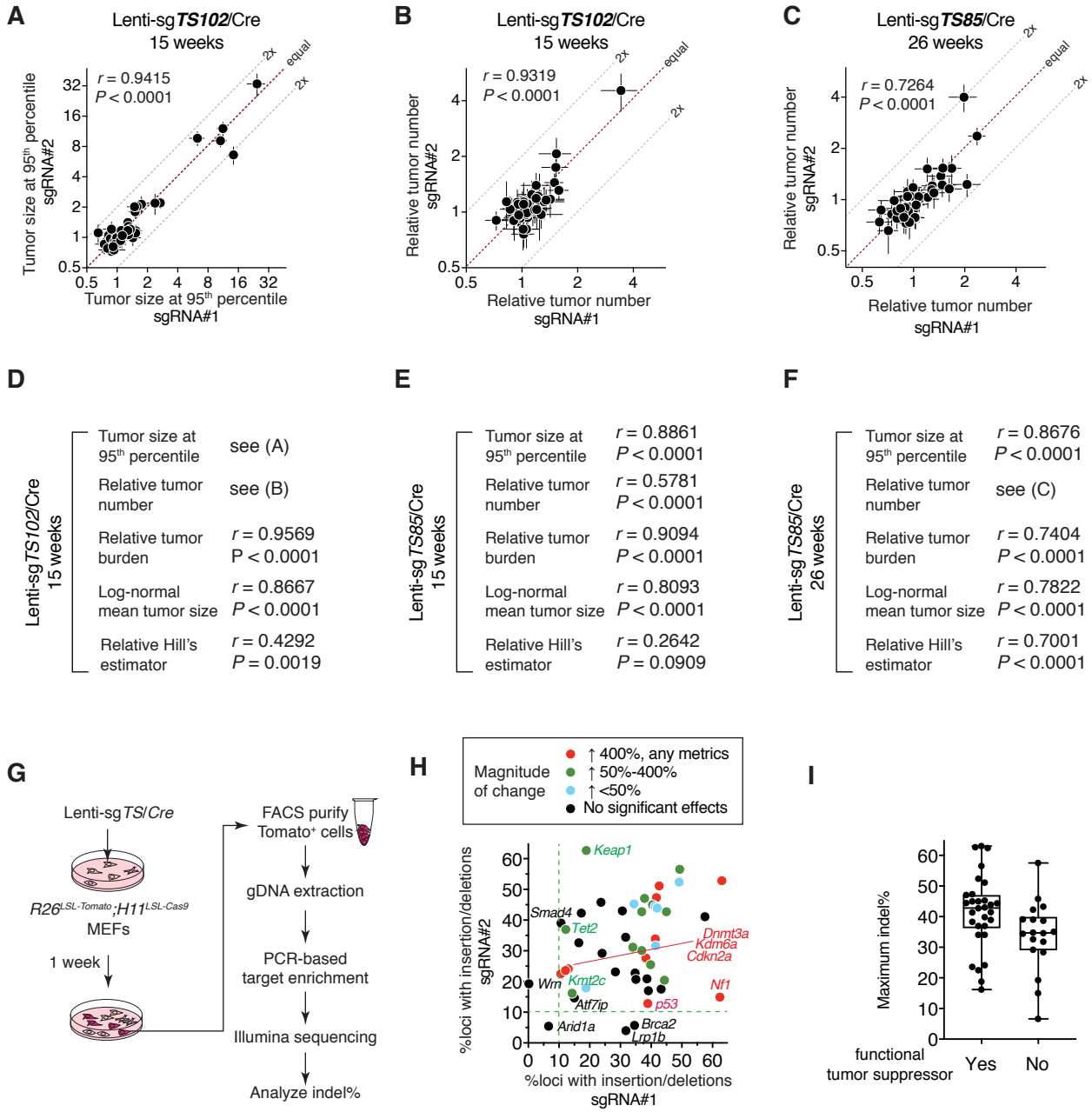
(A,B) The effect of *Kdm6a* inactivation on tumor size at 95th percentile is not significantly different between female and male mice. Volcano plots show all tumor suppressor targeting sgRNAs (highlighting *Kdm6a* targeting sgRNAs and inert sgRNAs). Each dot represents an sgRNA. Plots represent statistics calculated from aggregating all tumors from 47 (**a**) and 40 *KT;H11*$^{LSL\text{-}Cas9}$ mice (**b**).

(C,D) The effect of *Kdm6a* inactivation on relative tumor number is not significantly different between female and male mice. Volcano plots show all tumor suppressor targeting sgRNAs (highlighting *Kdm6a* targeting sgRNAs and inert sgRNAs). Each dot represents an sgRNA. Plots represent statistics calculated from aggregating all tumors from 47 (**c**) and 40 *KT;H11*$^{LSL\text{-}Cas9}$ mice (**d**).

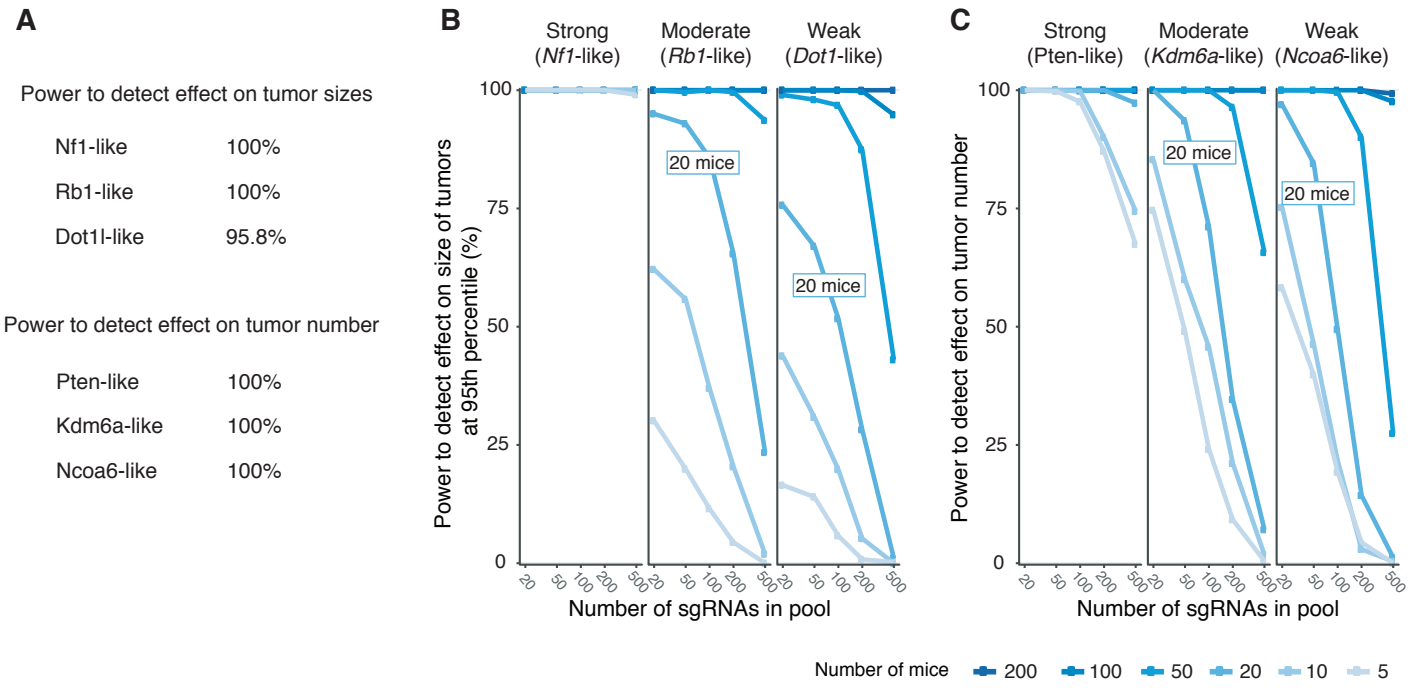
(E,F) The effect of *Kdm6a* inactivation on tumor initiation (relative tumor number) is consistent between female and male mice.

(G) The *Kdm6a* (also known as *Utx*) paralog *Uty* is expressed in cancer cells in autochthonous oncogenic *Kras*-driven lung tumors. Each dot represent an RNA-seq sample and the bar is the average (data are from cancer cells isolated from tumors in *Kras*$^{LSL\text{-}G12D}$*;p53*$^{f/f}$*;R26*$^{LSL\text{-}Tom}$ mice; Chuang *et al.*, Nature Medicine, 2017). Note that *Kdm6a* is expressed from the inactive X chromosome, thus explaining the higher expression of *Kdm6a* in females relative to males.

(H) Human lung adenocarcinomas with *KDM6A* mutations are not enriched in female patients, which would be expected if UTY could compensate for *KDM6A* inactivation in males.



Supplementary Figure 15. sgRNAs targeting the same genes elicit similar effects on tumor development.
(A-C) The tumor size at 95th percentile and relative tumor number of sgRNA#1 versus sgRNA#2 at the indicated time points. Pearson correlation coefficient (*r*) and *P*-value (two tailed) are indicated. Both sgRNAs targeting the same putative tumor suppressor elicit consistent and similar changes in the indicated metrics, substantiating the power to detect tumor suppressors and effectiveness of the sgRNA design parameters used.
(D-F) Pearson correlation coefficient (*r*) and *P*-value (two tailed) suggest that both sgRNAs targeting the same putative tumor suppressor elicit consistent and similar changes in size at 95th percentile, relative tumor burden, log-normal mean tumor size, relative tumor number and relative Hill's estimator, substantiating the effectiveness of the sgRNA design parameters used.
(G) Schematics of cutting efficiency analysis in cell lines. *R26*$^{LSL\text{-}Tomato}$*;H11*$^{LSL\text{-}Cas9}$ mouse embryonic fibroblasts (MEFs) were transduced with each individual Lenti-sg*TS*/Cre vector, purified by FACS for Tomato^{positive} cells after 1 week, and subject to analysis of insertion and deletion (indel) rates.
(H) The indel rates for sgRNA#1 versus #2 for each gene are plotted. Tumor suppressor genes are colored by their magnitude of effects. Red dots indicate tumor suppressor genes which increase any metric by >4 fold. Green dots indicate tumor suppressors with moderate effects. Blue dots indicate tumor suppressors with weak effect (<50% increase that is statistically significant). Dashed lines indicate the putative minimal insertion/deletion rate that enables detection of tumor suppressive effects (defined by sg*Cdkn2a*#1). Both sgRNAs targeting *Arid1a* are below this threshold, suggesting that we may miss *Arid1a* due to inefficient sgRNAs. *Wrn*, *Brca2* and *Lrp1b* each have only one sgRNA that is above this threshold.
(I) The maximal indel rate for sgRNAs targeting each gene. Boxes represent quartiles and median, while whiskers represent minimum and maximum. Most genes whose inactivation did not alter tumor growth have at least one sgRNA that generates indels comparible to those genes that do have effects.

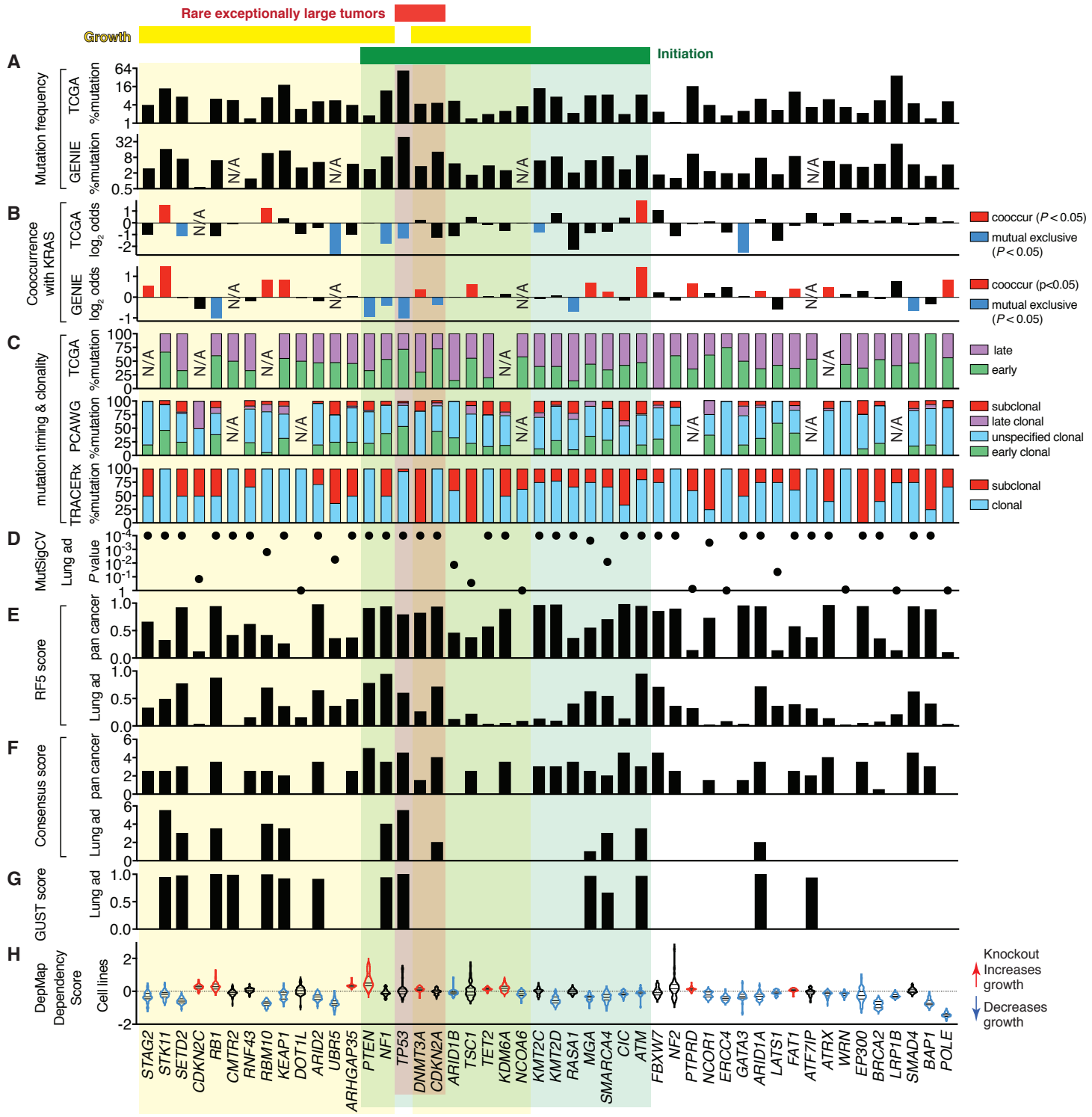


Supplementary Figure 16. Tuba-seq based metrics of tumor suppression are powered to uncover effects across a range of experimental parameters.

(A) The power to detect tumor suppressor genes in the current dataset. Tumor suppressors that exhibit strong, medium and weak effect in restraining tumor sizes are exemplified by Nf1, Rb1 and Dot1l (*Nf1*, sg#1; *Rb1*, sg#1; and *Dot1l,* sg#1), respectively. Tumor suppressors that exhibit strong, medium and weak effect in restraining tumor numbers are exemplified by Pten, Kdm6a and Ncoa6 (*Pten*, sg#2; *Kdm6a*, sg#2; and *Ncoa6,* sg#1), respectively. Simulations were done by 10,000 bootstrap resampling of tumors of the defined genotypes from the 47 *KT;H11*$^{LSL\text{-}Cas9}$ (initiated with lentiviral vectors at 30,000 ifu) and 12 *KT* (initiated with Lenti-sg*TS102*/Cre at 90,000 ifu) 15 week cohorts, and the percentage of times that we could successfully detect the tumor suppressor effect is indicated.

(B) Overlaid lines indicate smallest tested cohort size to achieve indicated power level. As expected, power increases with mouse number and decreases with the number of pooled sgRNAs. Strong effects on tumor size and number can be reliably detected even with very small cohorts and large numbers of pooled guides. Moderate and weak effects require 20-100 mice, depending on the number of pooled guides. Power to detect statistically significant differences in the size of tumors at 95th percentile initiated with sg*TS* relative to sg*Inert*. Power is shown for representative strong, moderate and weak tumor suppressors (*Nf1*, sg#1; *Rb1*, sg#1; and *Dot1l,* sg#1) as a function of the number of pooled sgRNAs and mouse cohort size.

(C) Power to detect statistically significant differences in the number of tumors initiated with sg*TS* relative to sg*Inert*s. Differences in the original proportion of sgRNAs in the lentiviral pool are calibrated by tumor numbers in *KT* mice. Power is shown for representative strong, moderate and weak suppressors of tumor initiation (*Pten*, sg#2; *Kdm6a*, sg#2; and *Ncoa6,* sg#1) as a function of the number of pooled sgRNAs and mouse cohort size.



Supplementary Figure 17. ***In silico*** **and** ***in vitro*** **analyses predict some, but not all functional tumor suppressors.**
Legend on next page

Supplementary Figure 17. *In silico* and *in vitro* analyses predict some, but not all functional tumor suppressors.
(A) Mutation frequencies of the 48 genes in TCGA (n = 566 cases) and GENIE (v7.0, n ≤ 8522 cases and varies by gene) lung adenocarcinoma cohorts. N/A: *CMTR2, UBR5, NCOA6* and *ATF7IP* are not profiled in any of the GENIE (v7.0) sequencing panels.
(B) Cooccurrence with oncogenic KRAS mutations. Log2 odds ratios are calculated from TCGA and GENIE (v7.0) lung adenocarcinoma cohorts. N/A: *CDKN2C* mutations were not detected in KRAS mutants in TCGA dataset, and *CMTR2, UBR5, NCOA6* and *ATF7IP* are not profiled in any of the GENIE(v7.0) sequencing panels.
(C) Mutation timing in the TCGA lung adenocarcinoma dataset is assessed by the occurrence of mutation before or after genome doubling (n = 507 cases; Cancer Genome Atlas Research Network, 2014 and McGranahan *et al*., 2015). Mutation timing and clonality in PCAWG lung adenocarcinoma dataset is assessed by the number of clones per cell (n = 36 cases; Gerstung *et al*., 2020). Mutation clonality in the TRACERx lung adenocarcinoma dataset is assessed by multi-region sequencing (n = 61 cases; Jamal-Hanjani *et al*., 2017). N/A: data for indicated gene is not available.
(D) MutSigCV score quantifies the probability of a gene being significantly mutated (Lawrence et al., 2013). Here we show the data reported in 2015 (Kumar *et al*., 2015).
(E) Scores quantifying the likelihood of each gene being a driver gene, across all cancers (pan-cancer) and in lung adenocarcinoma, as reported by Kumar *et al*., 2015 (prediction score by random forest classifier model; RF5 TSG score).
(F) Scores quantifying the likelihood of each gene being a driver gene, across all cancers (pan-cancer) and in lung adenocarcinoma, as reported by Bailey *et al*., 2018 (weighted consensus score across computational tools; TCGA driver consensus).
(G) Scores quantifying the likelihood of each gene being a driver gene in lung adenocarcinoma, as reported by Chandrashekar *et al*., 2020 (prediction score by random forest classifier model; GUST score).
(H) DepMap dependency score calculated from 85 lung adenocarcinoma cell lines. Up to 37 of these cell lines have dependency scores for indicated genes, and we further filter out cell lines that harbor damaging mutations in each indicated gene. Bars in the violin plots indicate median and quartiles. Genes that have positive scores for >75% cell lines are highlighted in red, while genes that have negative scores for >75% cell lines are in blue.

Supplementary Table 3. Estimated true positive rate for our study. We estimated the true positive rate using each of the four metrics for the three experiments. Column T shows the number of sgRNAs with significant tumor suppressor effect where the replicate sgRNA is also significant. Column F shows the number of sgRNAs with significant tumor suppressor effect where the replicate sgRNA is not significant. True positive rate (%) is calculated as the probability of using two guides to successfully identify the tumor suppressor effects for either guide for the metric in the experiment. A true positive rate of 80% is favorable in experimental designs. As shown in the table above, the true positive rate of our experiments is quite high.

Dataset	Metric	T	F	True positive rate (%)
Lenti-sg*TS102*/Cre 15-week	Size across percentiles	26	6	96.0
	Log-normal mean	30	5	97.7
	Relative tumor burden	34	9	95.3
	Relative tumor number	26	12	89.7
Lenti-sg*TS85*/Cre 15-week	Size across percentiles	4	2	87.2
	Log-normal mean	6	4	83.3
	Relative tumor burden	2	7	43.7
	Relative tumor number	4	6	65.1
Lenti-sg*TS85*/Cre 26-week	Size across percentiles	10	5	88.2
	Log-normal mean	16	10	84.9
	Relative tumor burden	10	11	72.7
	Relative tumor number	22	4	97.2